

DOCUMENT RESUME

ED 130 623

IR 004 154

AUTHOR Clifton, Charles, Jr.; And Others
TITLE Grammatical Relations as Determinants of Sentence Similarity. Studies in Verbal Behavior Report Number 11.
INSTITUTION Minnesota Univ., Minneapolis. Dept. of Psychology.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 1 May 64
GRANT NSF-G-18690
NOTE 21p.; Archival document

EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.
DESCRIPTORS *Kernel Sentences; Negative Forms (Language); *Sentence Structure; *Transformation Generative Grammar

ABSTRACT

Generalized recognition of the kernel, passive, negative, and passive-negative transformations of a number of sentences was investigated. A significant amount of generalization was obtained within a transformationally-defined sentence family relative to the generalization obtained between sentence families when subjects were requested to register recognition of sentences previously read. A distance measure was used to compare generalization decrements between various constructions. The distance between sentences related by the passive transformation was less than between those related by the negative or a combination of negative and passive transformations. The distance between sentences related by negative transformations was not consistently less than that related by the combination of passive and negative. When the order of transformation was considered, the distance between sentences related by both passive and negative transformations equaled the sum of the distance between sentences related by the passive and negative. A satisfactory spatial representation of this data was Attneave's "city block" model. (KB)

* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *

ED130623

University of Minnesota
Department of Psychology

GRAMMATICAL RELATIONS AS DETERMINANTS
OF SENTENCE SIMILARITY

by

Charles Clifton Jr., Ida Kurcz

James J. Jenkins

STUDIES IN VERBAL BEHAVIOR

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

REPORT NUMBER 11

1 MAY, 1964

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

Supported by the National Science Foundation

Under Grant G - 18690

Principal Investigator: James J. Jenkins

Grammatical Relations as Determinants of Sentence Similarity¹

Charles Clifton, Jr., Ida Kurcz, and James J. Jenkins

University of Minnesota

Chomsky (1957) has argued that an elegant formal grammar of English can be constructed using transformation rules. These rules are applied to terminal strings, which are the outputs of a phrase structure grammar. Transformation rules specify that a particular structural change is to be made in a terminal string (or transformed terminal string) having a particular structural description.

For the present purposes, transformation rules can be considered as being applied to sentences, which are, in the grammar, the products of applying morphophonemic rules to the already transformed terminal strings. We can speak of applying the passive transformation to a kernel (simple, active, declarative, affirmative) sentence such as The man closed the box to produce a passive sentence, The box was closed by the man. Similarly, the negative transformation can be applied to a kernel sentence to produce a negative sentence, The man did not close the box. (In this particular example, Chomsky's tentative grammar specifies that two transformations - the negative transformation and the 'do' transformation - be applied. Since these two transformations are not independent, the two of them are here considered as a single transformation.) Applying both the passive and the negative transformations to the kernel (or applying the negative transformation to the passive) produces a passive-negative sentence, The box was not closed by the man.

Although the operation is not formally in the grammar, we will here speak also of 'removing' a transformation.² For instance, removing the passive transformation from a passive sentence would produce a kernel.

The set of sentences which can be produced from a given kernel by applying each

transformation and each permissible combination of transformations can be referred to as a sentence family. Internal relationships within such a family are completely described by the transformational relationships among the family members.

Miller (1962) has suggested that Chomsky's formal analysis provides a promising tool for a psychological grammar, and has hypothesized that transformations may be coordinated with psychological processes, or at least that behavioral relationships among sentences will prove to be a systematic function of the transformational relationships. He has provided relevant data on sentence identification times, indicating that the time taken to match two sentences, of different constructions but based on the same kernel, is a function of the number of transformation rules which must be applied or removed to transform one sentence to the other.

The present study reports another behavioral relationship among sentences of different syntactic constructions, namely, the degree to which a motor response generalizes among the sentences. A technique developed by Mink (1963) for studying generalization among associatively related words was used. In the present study, generalization among the kernel (K), passive (P), negative (N), and passive-negative (PN) constructions of a sentence was measured. It was thought that the data might provide some evidence to support or refute the hypothesis that the transformational relationships provide a useful description of the behavioral relationships among the sentences.

Method

Materials

Two sets (A and B) of 12 past-tense kernel sentences (such as The man closed the box, John hit the ball, and The horse ate the hay) were constructed. The passive, negative, and passive-negative transform of each sentence was generated. Each set was divided into four subsets (1, 2, 3, 4) of three sentences each. Eight 12 sentence training list forms were constructed. Each training list form was composed of all the

sentences from one set (A or B), the three sentences from one subset being in the kernel construction, the three from another subset in the passive construction, the three from still another subset in the negative construction, and the three from the final subset in the passive-negative construction. The composition of the A training list forms is indicated in Table 1. The composition of the B training list forms was identical, except that sentences from set B were substituted for sentences from

Insert Table 1 about here

set A. Four training list randomizations were made, the same randomizations being applied to each training list form.

The test list contained 96 sentences, consisting of the four constructions of each of the 24 basic sentences used in the experiment. Thus, there were, on the test list, 48 sentences unrelated to sentences which were on a particular training list (new sentences), and 48 sentences which were transformationally related to sentences on the given training list. For instance, among these latter, there were 12 test list kernels related to training list sentences: three identical with sentences which appeared in the kernel construction on the training list, three related to training list passives, three related to training list negatives, and three related to training list passive-negatives. The first three of these sentences will be referred to as 'training list kernels,' the second three as 'kernels of training list passives,' the third three as 'kernels of training list negatives' and the fourth as 'kernels of training list passive-negatives. The same notational scheme will be followed for the test list sentences of the remaining constructions.

Each such set of three sentences, defined by the test list construction of the sentences and the construction of the training list sentences to which they were related, will be referred to as a test sentence type.

Apparatus

A Lafayette memory drum was used to present the sentences to the Ss. A telegraph key was mounted in front of the memory drum and connected to a 12-volt transformer-operated buzzer audible to both S and E. An Ampex tape recorder was used to present the instructions.

Subjects and Procedures

Thirty-two undergraduates at the University of Minnesota were used as Ss. Each S heard instructions indicating that he was going to see a list of sentences and instructing him to press the telegraph key immediately after he silently read each sentence. S was instructed to try to remember the sentences on the list so that he could recognize them later. The four randomizations of the training list assigned to the S were then presented on the memory drum at the rate of one sentence every four seconds with a four second interval between randomizations. Immediately after the presentation of the fourth randomization S heard instructions indicating that he was going to see a longer list, containing the sentences he had seen on the training list as well as some others. He was instructed to press the telegraph key whenever he thought he recognized a sentence, but not to press it when he was sure he did not recognize a sentence. He was then given one presentation of one of four different randomizations of the test list.

- Subjects were assigned consecutively to training list forms in the order A-I,..., A-IV, B-I,..., B-IV on the basis of their order of appearance. Four Ss were assigned to each training list form. One S in each form received each test list randomization. Thus, the design may be seen to consist of four randomized replications of a single counterbalanced design.

Results and Discussion

Table 2 presents the data on test list key pressing in terms of the proportion

Insert Table 2 about here

of opportunities on which presses were actually made. The number of presses to each test sentence type (omitting the types composed of sentences identical to training list sentences) was compared with the number of presses to the new (control) sentences of the same construction on the test list. In all 12 of the possible comparisons, significantly more ($p < .01$, one-tail Wilcoxon test) responses were made to the transforms of training list sentences than to the appropriate new sentences, after correcting for the fact that there were four times as many opportunities to press to new sentences of a given construction than to generalization sentences of a given test sentence type. It is obvious that an impressive amount of generalization was obtained within a transformationally-defined sentence family, relative to the generalization obtained between sentence families.

The number of responses made to training sentences and transforms of training sentences was further analyzed. Each test sentence can be placed in one of four categories on the basis of its transformational relationship to a training list sentence. These four types of transformational relationships, and the test sentence types falling into the category defined by each relationship, are:

1. Identity

training kernel	(TK)
training passive	(TP)
training negative	(TN)
training passive-negative	(TPN)

2. Passive transformation

passive of training kernel	(P of TK)
kernel of training passive	(K of TP)
passive-negative of training negative	(PN of TN)
negative of training passive-negative	(N of TPN)

3. Negative transformation

negative of training kernel	(N of TK)
kernel of training negative	(K of TN)
passive-negative of training passive	(PN of TP)
passive of training passive-negative	(P of TPN)

4. Passive plus negative transformations

passive-negative of training kernel	(PN of TK)
kernel of training passive-negative	(K of TPN)
negative of training passive	(N of TP)
passive of training negative	(P of TN)

The test list pressing data can be grouped on the basis of these transformation relationships to training list sentences, and are so grouped in Table 2. Within each category fall sentences of the four different test list constructions. These four different sentence constructions were considered as providing repeated measures within each category. A transformational relationship X Ss, repeated measures, analysis of variance was performed on these data. Significant variation among the means of the four categories was found ($F = 24.408$, $df = 3/93$, $p < .001$). The mean of each category was significantly different from the mean of each other category (t - tests $p < .01$ for each comparison).

It is apparent that a significant generalization decrement from a training sentence to its transforms was obtained, and that the magnitude of this decrement depended on the particular transformation which related the training sentence to its transformed test sentence. It was thought that these decrements would be best described by constructing a tentative scale of distance between the various syntactic constructions, with distance related to generalization decrement.

The mean (over Ss) number of responses made to sentences of a given test sentence type can be viewed as a measure of the proximity of the two sentence constructions

defining the test sentence type - the test sentence construction and the construction of the training sentence to which it was related. It should be noted that this proximity measure is directional. A reciprocal measure of the proximity between the two constructions is determined by the pressing to sentences of the test sentence type defined by the same constructions, but with the test and training constructions reversed. For instance, the number of presses to negatives of training kernels determines the proximity of negative to kernel, while the number of presses to kernel of training negatives determines the proximity of kernel to negative, a proximity which is reciprocal to the first.

Such a set of proximities can be related by a monotonic negative function to a set of distance measures (D's.) A wide variety of such functions are possible and defensible, but, as a first approximation, the simple linear function, $D = 2.37 - \text{proximity}$, was chosen. The additive constant 2.37 equals the mean level of test list responding to all training list sentences, and was chosen to make the mean distance from any construction to itself equal to zero.³ These D values, together with the means of the reciprocal D values, are indicated in Figure 1, which is not to be taken as a spatial model of the syntactic generalization space, but simply as a convenient way of displaying the D values.

Insert Figure 1 about here

First, a test sentence type X Ss analysis of variance was performed on the D values of the transforms of training sentences (that is, the test sentence types with the types composed of training sentences omitted.) (This analysis, and the following analyses, were actually performed on the proximity values. However, due to the linear relation between proximity and D, these are equivalent to analyses of the D values.) Significant variation was found among the means of the 12 sentence types ($F = 10.773$, $df = 11/341$, $p < .001$). In contrasting the reciprocal pairs of D's - e.g., contrasting the D for negatives of training kernels with the D for kernels of training

negatives - no significant differences were found. The t 's for the six such contrasts ranged from 0.587 to 1.419 with 341 df . This indicates that generalization is not significantly asymmetrical, that is, for present purposes reciprocal D values can be considered equal.

Because of this apparent symmetry, the data from reciprocal distances were combined, and are reported as the means in Figure 1. The classification based on the 12 test sentence types used in the immediately preceding analysis of variance was replaced by a classification based on the six transformational generalization classes which remained after the pairs of reciprocal D's were classed together, and the analysis was repeated, with the two D values constituting each pair of reciprocal D's being treated as repeated measures. A significant effect of transformational generalization class was found ($F = 22.866$, $df = 5/155$, $p \leq .001$).

The Duncan Multiple Range Test was used to make individual comparisons of the mean D values associated with the transformational generalization classes. Results of the test are reported in the ordering of the classes appearing at the end of this paragraph. Here, the transformational generalization classes are designated by the training and test sentence constructions of their member sentences (e.g., the class consisting of negatives of training passive-negatives and passive-negatives of training negatives is designated N-PN), and are ordered in terms of increasing D values. The D for each class is indicated immediately below the class designation. The result of the range test is indicated by the lines above the classes. The D values for classes falling under the same line were not significantly different from each other at the .05 level. The D values for classes not falling under the same line were significantly different from each other at the .05 level.

N-PN K-P P-PN			P-N K-N K-PN		
(.51)	(.71)	(.98)	(1.46)	(1.59)	(1.76)

It appears that sentences related by the passive transformation are closer together by the present distance measure than are sentences related by the negative transformation or by the combination of passive and negative transformations. This conclusion is supported by the first analysis of variance reported (the transformational relationship X Ss anova). However, sentences related by the negative transformation are not consistently closer together than sentences related by the combination of the passive and negative transformations (although the first analysis of variance reported does indicate that there is an overall difference between sentences related by the negative transformation and sentences related by the combination of passive and negative transformations). The hypothesis that the distance between sentences is directly related to the transformational difference between them would seem to predict that all sentences differing by the combination of two transformations would be further apart than sentences which differ by only one of the two transformations. Specifically, it would suggest that the distance between kernel and negative should be less than the distance between passive and negative.

A possible resolution is suggested by the observation that the negative transformation produces a greater generalization decrement (gives rise to a greater D value) when it is applied to active sentences than when applied to passive sentences. The passive and passive-negative constructions appear to be closer together than the kernel and negative constructions.

Referring to figure 1, it is noted that the distance between kernel and passive-negative (1.76) is close to the sum of the kernel - passive and passive-passive-negative distances (1.69), while it is somewhat less than the sum of the kernel - negative and negative - passive-negative distances (2.10). Also, the distance between passive and negative (1.46) is approximately equal to the sum of the passive - passive-negative and passive - negative-negative distances (1.49), but noticeably less than the sum of the passive - kernel and kernel - negative distances (2.30). This kind of

evidence -- the apparent additivity of behavioral effects of transformations - has been taken (Miller, 1962) as a strong point in favor of the transformational interpretation of behavioral sentence relationships. Further, the necessity for considering the order of application or removal of transformations, implied by the present data, combined with the previous observation that the effect of the negative transformation depends on the form of the sentence to which it is applied, may provide an explanation for the lack of a consistent difference between the effect of applying the negative transformation and the effect of applying both the negative and passive transformations.

It seems somewhat premature to attempt to construct a spatial representation of the set of distances. In view of the potential helpfulness of such representation in portraying the data, however, certain statements about the possible spatial models should be made.

It is apparent, from the data on the 'diagonal' distances presented immediately above, that the distances cannot be represented using any two-dimensional Euclidian model. It would be possible to use a model such as that implied by Miller (1962), in which one defines only points (corresponding to syntactic constructions) and paths (lines between constructions which differ from one another by only a single transformation), and in which distance between two points is coordinated with the length of the path connecting the two points (or the sum of the lengths of the shortest set of connected paths connecting the points, if the points correspond to constructions differing by more than a single transformation.) Such a model is consonant with the data. This type of model also fits the data presented by Miller on sentence 'transformation time,' although, unlike the present data, the time data indicated that the distance between sentences related by the negative transformation is less than the distance between sentences related by the passive transformation. The fit of such

a model to the data can be taken to support the hypothesis that relationships among sentence constructions can be described simply by their transformational relationships.

An apparently satisfactory spatial representation of the present data can be based on Attneave's (1950) 'city block' model. The city block model implies that the syntactic constructions can be arranged in a set of dimensions, and that the distance between any two constructions is equal to the sum, over all dimensions, of the differences between the projections of the two constructions on each dimension. The present data can be represented in two dimensions (for an example of one such representation, see Figure 2), which could be tentatively titled 'passivity' and 'negativity.'

Insert Figure 2 about here

However, not too much psychological sense can be made of a syntactic interpretation of these dimensions at the present time, since, for instance, either the PN construction must be represented as being less passive than the P construction, or the N construction as more passive than the K construction. Further, four points are hardly enough to test the applicability of the model to sentence generalization data. It would seem wise to hold further evaluation of the applicability of the city block model in abeyance.

The three most obvious bases on which the generalization data can be explained can be termed 'formal similarity of the sentences,' 'semantic similarity of the sentences,' and 'syntactic (transformational) relationships of the sentences.' It would be difficult to account for the data by reference to formal (physical) similarities between the sentences of different constructions. While no measure of formal similarity is available, it seems reasonable to say that the similarity between the configurations of kernel and passive sentences is less than that between the configurations of kernel and negative sentences. There was, however, more generalization in the former case than in the later.

It is not obvious that syntactic differences between sentence constructions can be considered independently of differences in meaning. These differences - syntactic and semantic - are certainly not separated in the constructions used in the present experiment. Sentences related by the negative transformation obviously differ in meaning, and sentences related by the passive transformation can be considered as differing in connotative meaning, at least. It can be argued that the fit of the data to the Miller type model, with its exclusive dependency on transformations for describing relationships among sentences, supports a syntactic interpretation of the data. Conversely, it can be argued that the dimensions in the equally satisfactory block model can most easily be coordinated with dimensions of meaning. These arguments are, however, less than compelling. It seems likely that the decision between the two interpretations must wait a more thorough explication of psychological mechanisms underlying the phenomena of grammar as well as the development of a more comprehensive theory of meaning than is now available.

Summary

Generalization of a motor response among the kernel, passive, negative, and passive-negative forms of a number of sentences was investigated. The generalization among such constructions, relative to generalization among unrelated sentences, was found to be highly significant. A distance metric was used to compare the generalization decrements obtaining between the various constructions. It was found that the distance between sentences related by the passive transformation was less than the distance between sentences related by the negative transformation or by the combination of the negative and passive transformations. The distance between sentences related by the negative transformation was not consistently less than the distance between sentences related by the combination of passive and negative transformations.

However, when the possible orders of application of the transformations were taken into account it appeared that the distance between sentences related by the combination of passive and negative transformations equalled the sum of the distance between sentences related by the passive transformation and the distance between sentences related by the negative transformation. Models for representing the distance data were considered.

References

- Attneave, F. Dimensions of similarity. Amer. J. Psychol., 1950, 63, 516-556.
- Chomsky, N. Syntactic Structures. The Hague, The Netherlands: Mouton and Co., 1957.
- Miller, G. A. Some psychological studies of grammar. Amer. Psychologist, 1962, 17, 748-762.
- Mink, W. D. Semantic generalization as related to word association. Psychol. Rep., 1963, 12, 59-67.

Footnotes

1. This research was supported in part by grant G - 18690 from the National Science Foundation, J. J. Jenkins, principal investigator. Mr. Clifton held a National Science Foundation Cooperative Graduate Fellowship at the time the reported research was carried out. Dr. Kurcz is a member of the faculty of the University of Warsaw.

2. By 'removing a transformation' we simply mean that the construction (e.g., passive) from which the transformation (e.g., passive transformation) is removed to produce a second construction (here, kernel) can be produced by the application of the transformation to the second construction. We do not propose to introduce a new operation corresponding to 'removing a transformation.'

3. It might be argued that a different additive constant should be used in the calculation of the D's for the test list sentences of the different constructions. This would allow a correction for any possible responding bias favoring some constructions over others (such a bias, in fact, apparently occurring in the present data.) It might even be argued that such a response bias would make a linear function suspect, and that the function should be such as to allow, say, a multiplicative correction for the bias. Further data may support the wisdom of such a position. For the present, however, the simplicity and orderliness of the results obtained when no such correction is made oppose such an argument. Further, the occurrence of apparent 'response bias' - such as a higher level of responding to test passives - can be derived from the pattern of distances among the constructions as obtained using the simple, linear function.

Table 1

Construction of Form A Training Lists

(Entries indicate the construction in which sentences of a particular sentence subset appear in a particular training list form.)

Sentence subset (three sentences in each subset)	Training List Form			
	A-I	A-II	A-III	A-IV
A1	K	P	N	P-N
A2	P	N	P-N	K
A3	N	P-N	K	P
A4	P-N	K	P	N

Note. -- Construction of Form B lists is similar, with set B sentences replacing set A sentences.

Table 2

Proportion of Possible Presses to Test List Sentence Types

(Letters in body of table indicate test sentence type)

Transformational relationship between test and training sentence					
Test List Sentence	Identity	Passive	Negative	Passive + negative	Unrelated (control)
Construction	Identity	Passive	Negative	Passive + negative	Unrelated (control)
Kernel	.75	.53	.24	.17	.003
	(TK)	(K of TP)	(K of TN)	(K of TPN)	
Passive	.87	.57	.46	.35	.003
	(TP)	(P of TK)	(P of TPN)	(P of TN)	
Negative	.70	.58	.28	.25	.003
	(TN)	(N of TPN)	(N of TK)	(N of TP)	
Passive Negative	.85	.66	.47	.24	.003
	(TPN)	(PN of TN)	(PN of TP)	(PN of TK)	



